Journal of Bioscience and Bioengineering
VOL. 114 No. 2, 232–236, 2012

www.elsevier.com/locate/jbiosc

# Gas chromatography–mass spectrometry based metabolic profiling for the identification of discrimination markers of *Angelicae Radix* and its application to gas chromatography–flame ionization detector system

Shizu Kobayashi,[1] Sastia Prama Putri,[1] Yutaka Yamamoto,[2] Kang Donghyo,[2] Takeshi Bamba,[1] and Eiichiro Fukusaki[1,*]

*Department of Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan*[1] and *Tochimoto Tenkaido Co., Ltd, 3-21 Suehiro-cho, Kita-ku, Osaka 530-0053, Japan*[2]

Received 4 February 2012; accepted 12 March 2012
Available online 25 May 2012

**Gas chromatography (GC)-based metabolomics technologies were applied for quality control of *Angelicae Radix*, an herbal medicine commonly used in Japan and China. Since *Angelica* roots are priced and graded differently based on their species and cultivation area, there is a need for a simple and reproducible method to discriminate *Angelica* roots. Here, we used GC–MS profiling data to construct a discrimination method for species and cultivation area of *A. Radix*. Seventy-six primary metabolites were identified. The quality factors of *A. Radix* were successfully classified using metabolic profiling and the orthogonal projections to latent structures-discriminant analysis (OPLS-DA) technique. Sorbitol and a glucose/4-aminobutyric acid combination were chosen as bio-markers from S-plot of OPLS-DA. Application of these selected bio-markers to a more practical and cost-efficient system, namely gas chromatography–flame ionization detector (GC–FID) system were also assessed. As a result, the same separations of sorbitol, glucose and 4-aminobutyric acid in box plots were obtained from GC–FID data. Our results demonstrate that GC-based metabolic markers can be readily applied for the establishment of a practical quality control method for *A. Radix*.**


[**Key words:** *Angelicae Radix*; Gas chromatography–mass spectrometry (GC–MS); Gas chromatography–flame ionization detector (GC–FID); Metabolomics; Quality control]

The dry roots of *Angelica* (*Angelicae Radix*) have been used as a traditional crude drug for a wide variety of gynecological disorders in eastern Asia. Japanese Pharmacopeia 16th (JP) approves the use of *Angelica acutiloba* Kitagawa (Yamato-toki in Japanese) and *A. acutiloba* Kitagawa var. *sugiyamae* Hikino (Hokkai-toki in Japanese), whereas *A. sinensis* is the widely used species in China. Recently, the supply of *A. Radix* produced in Japan has been insufficient, meeting only approximately 44% of the domestic demand (1). To address this shortage, *A. acutiloba* that was initially exported to China and subsequently cultivated in China is now imported to and circulated within Japan. In Japan, *A. acutiloba* that is produced domestically is generally retailed at a higher price than *A. acutiloba* produced in China or *A. acutiloba* var. *sugiyamae* (2). Thus, discrimination of *A. Radix* based on their species and cultivation area is becoming an important issue in herbal medicine industry.

Species discrimination has been done using genetic analyses as well as metabolome analyses. Species of *A. Radix* could be discriminated using DNA sequence analysis of the 5S-rRNA spacer domains and the plastid genome (3). Random Amplification of Polymorphic DNA (RAPD) has also been used to discriminate *A. acutiloba*, *A. acutiloba* var. *sugiyamae*, and *A. sinensis* (4,5). Although DNA is an extremely stable macromolecule that is not affected by external factors (6), the potential use of genetic based method is hindered by several technical factors, such as contaminations of non-target DNA in PCR samples (7) or low reproducibility in RAPD method (8). Discrimination of *A. acutiloba* and *A. sinensis* based on their chemical constituents has been mainly focused on the major chemical constituents that are directly related to the pharmacological activity of *A. Radix*, such as ligustilide and xanthotoxin (2). Studies on the identification of species based on the non-polar compound content (9,10), and by volatile compounds, such as ligustilide and butyliden pthalide (11) have also been reported.

Although the information on pharmacological compounds holds significant importance on the study of *Angelica*, the quality of *A. Radix* is currently determined based on their sensory quality such the appearance, aroma and taste instead of solely based on their active components. Hence, the use of pharmacological compounds as markers appear to be inadequate as it is not in line with the method used for grading the quality of crude samples in commercial markets. Sweetness of sensory qualities for *A. acutiloba* is used as an important quality factor since 1783 (1) at least. Moreover, sugars have been described as the compounds that are highly

* Corresponding author. Tel./fax: +81 6 6879 7424.
*E-mail address:* fukusaki@bio.eng.osaka-u.ac.jp (E. Fukusaki).

1389-1723/$ – see front matter © 2012, The Society for Biotechnology, Japan. All rights reserved.
doi:10.1016/j.jbiosc.2012.03.022

affected by processing methods (12). Therefore, development of methods for quality evaluation and discrimination of final products of *A. Radix* based on primary metabolites are considered more applicable for use in the industry.

Our group has recently reported several metabolomics studies for the discrimination of *A. acutiloba* and *A. acutiloba* var. *sugiyamae* using nuclear magnetic resonance (NMR) (13), gas chromatography–mass spectrometry (GC–MS) (14), pyrolysis GC–MS (15), and ultra-performance liquid chromatography–MS (UPLC–MS) (16). However, the discrimination was performed only by qualitative analysis, and the statistical significance of candidate markers was not mentioned in the reports. One recent report that used statistic analysis for the evaluation of markers by NMR and UPLC–MS is the study on *A. gigas* from three different regions in Korea (17). The samples analyzed only represented cultivation area as quality factor. To evaluate the quality of *A. Radix* in the global market, a method to assess markers for both species and cultivation area using a more quantitative analysis with statistical significance by *t*-test is needed. At the same time, the established method should also be reliable, rapid, and cost-effective for real application in industry.

In this study, we aim to establish a practical method for the discrimination of species and geographical origin of *A. Radix*. To achieve our goals, we utilized GC–MS profiling data or primary metabolites to identify statistically significant markers for the discrimination of *A. Radix*. GC–MS was initially used for the purpose of compound identification with high accuracy, sensitivity and reproducibility. Subsequently, we showed an application of the identified markers to a more practical and cost-effective instrument, gas chromatography with flame ionization detector (GC–FID) (18). GC–FID is suitable for routine quality control of *A. Radix* because it is a relatively inexpensive instrument, is currently available in most of small factories and have already been recognized in the JP. We focused on effective methods of classifying *A. Radix* during industrial control rather than on the pharmaceutical effects of the samples.

As an appropriate tool to identify markers for classification, Orthogonal projections to latent structures-discriminant analysis (OPLS-DA) models were applied. The variance of these compounds between sample groups was statistically evaluated by *t*-test. We further assess the application of these markers using a simpler and cost-effective instrument, GC–FID. The GC–FID usability was also shown from the standpoint of the statistical significances in *t*-test and peak intensities by box plots. Our results demonstrate that GC-based metabolomics technology is a powerful tool for the establishment of a practical quality control method for the discrimination of species and cultivation area of *A. Radix*.

## MATERIALS AND METHODS

**Samples** All the samples of *A. Radix* were provided from Tochimoto Tenkaido Co. Ltd. (Osaka, Japan). Species identifications of the samples were performed by description of a crude drug and microscopic identification by trained specialists prior to analysis of this study. Sample details are shown in Table 1. Nara, Hokkaido and Gansu are known as main cultivation areas of *A. acutiloba*, *A. acutiloba* Kitagawa var. *sugiyamae* Hikino (1) and *A. sinensis* (19), respectively. The specimens were first sliced or chipped before they were ground into a fine powder. The samples were kept in 50 ml plastic tube with light shielding in −30°C freezer before use. Sample extraction and analysis conditions for GC-based metabolic profiling was carried out as described previously (14) with some modifications which could be found in Supporting Information. Details for sample extraction, analysis conditions for GC–MS and GC–FID, data preprocessing and multivariate analysis are also available as Supporting Information.

**Chemicals** For GC–FID analyses of markers, sorbitol, 4-aminobutyric acid were purchased from Wako Pure Chemical Industries Ltd. (Osaka, Japan). Glucose was purchased from Nacalai Tesque (Kyoto, Japan).

**TABLE 1.** *Angelicae Radix* samples used in this study.

| Population number | Group number | Collection year | Cultivation region | Species |
|---|---|---|---|---|
| 1 | 1 | 2008 | Nara, Japan | *A. acutiloba* |
| 2 | 1 | 2008 | Nara, Japan | *A. acutiloba* |
| 3 | 1 | 2008 | Nara, Japan | *A. acutiloba* |
| 4 | 1 | 2007 | Nara, Japan | *A. acutiloba* |
| 5 | 1 | 2006 | Nara, Japan | *A. acutiloba* |
| 6 | 1 | 2005 | Nara, Japan | *A. acutiloba* |
| 7 | 1 | 2007 | Tokushima, Japan | *A. acutiloba* |
| 8 | 2 | 2008 | Zhejiang, China | *A. acutiloba* |
| 9 | 2 | 2007 | Zhejiang, China | *A. acutiloba* |
| 10 | 2 | 2005 | Zhejiang, China | *A. acutiloba* |
| 11 | 2 | 2004 | Zhejiang, China | *A. acutiloba* |
| 12 | 2 | 2004 | Zhejiang, China | *A. acutiloba* |
| 13 | 2 | 2004 | Zhejiang, China | *A. acutiloba* |
| 14 | 3 | 2007 | Hokkaido, Japan | *A. acutiloba* Kitagawa var. *sugiyamae* Hikino |
| 15 | 3 | 2007 | Hokkaido, Japan | *A. acutiloba* Kitagawa var. *sugiyamae* Hikino |
| 16 | 3 | 2006 | Hokkaido, Japan | *A. acutiloba* Kitagawa var. *sugiyamae* Hikino |
| 17 | 4 | 1995 | Gansu, China | *A. sinensis* |
| 18 | 4 | 2008 | Gansu, China | *A. sinensis* |
| 19 | 4 | 2008 | Gansu, China | *A. sinensis* |
| 20 | 4 | 2008 | Gansu, China | *A. sinensis* |
| 21 | 4 | 2008 | Gansu, China | *A. sinensis* |

## RESULTS AND DISCUSSION

**GC–MS analysis** GC–MS was used to identify and quantify low-molecular-weight hydrophilic compounds in *A. Radix* samples. After performing data processing and peak determination as described in Materials and Methods, we were able to identify seventy-six compounds in these samples (Table S1). The total number of identified compounds was much higher than 22 metabolites in our previous study (14), mostly due to the operation of AI output, a newly developed GC–MS peak annotation system (20).

**Principal component analysis** Principal component analysis (PCA) was used for exploring the data structure obtained by GC–MS. As shown in Fig. S1, the first and second principle components (PC) of the PCA score plot represented 38.5% and 35.2% of the total variance of the samples, respectively. The variance of the species differences was successfully captured in PC2. The loading plot of PC2 revealed that proline contributed to the discrimination of *A. sinensis* from other species. This result was consistent with a previous report stating that accumulation of proline, which can range up to 100-fold, varies in plants (21). Therefore, proline could become a valuable marker for distinguishing *A. Radix*.

**Orthogonal projections to latent structures-discriminant analysis (OPLS-DA) classification** We then further analyzed the components that contributed to the separation of *A. Radix* based on species and agricultural origin using orthogonal projections to latent structures-discriminant analysis (OPLS-DA) and component profile data generated by GC–MS. S-plot (22) of OPLS-DA was used because it helps identifying statistically significant metabolites, based both on contributions to the model and their reliability. A number of seemingly different data-analytical problems can be expressed as regression problems with a special coding of Y or X in linear discriminant analysis (LDA) and analysis of variance (ANOVA) (23). For example, there were problems such as dimensionality, multicollinearity, noise and missing data (24). With many and collinear variables, a PLSR solution can therefore be formulated for each of these. In addition, advantage of OPLS compared to PLS is that the model is rotated so that class

separation is found in the first predictive component (22). The first component of OPLS makes model interpretation easier. S-plot of OPLS-DA was used to search for reliable bio-markers because S-plot was an excellent tool for interpretation of OPLS-DA. The samples were divided into four groups based on species and cultivation area, *A. acutiloba* from Japan, *A. acutiloba* from China, *A. acutiloba* var. *sugiyamae* and *A. sinensis* (Table 1). Phylogenetic analysis of *A. Radix* by RAPD showed that *A. sinensis* could be completely discriminated from the other two *Angelica* species (3,4). Thus, we started classifying all *A. Radix* samples by distinguishing *A. sinensis* from the others, and subsequently separated *A. acutiloba* and *A. acutiloba* var. *sugiyamae*. Discrimination of *A. acutiloba* based on its agricultural origin was also performed. All OPLS-DA models were constructed with the number of latent variable of 1 and $Q^2$ greater than 0.8 (Table S2). $Q^2$ is the cross-validated predictive ability (22). Whether or not a prediction model is good can be judged in terms of the $Q^2$ value, and a model is considered to be good if $Q^2 > 0.5$; if $Q^2 > 0.9$, a model is considered to have an excellent predictive ability (24).

Additionally, permutations test were performed in the PLS-DA model to validate each OPLS-DA model according to validation methods in previous work of *A. gigas* (17). All $Q^2$ and $R^2$ values were higher in the permutation test than in the real model, revealing great predictability and goodness of fit. External validation aims to address the accuracy of a model in samples from different samples. For the prediction, the one or two samples (a test data set) were left randomly from each region (a training data set) and the OPLS-DA prediction model was performed three times without them. The cutoff of the prediction was 0.5. As a result, almost all samples were correctly classified (Fig. S2). In Table S3, $R^2$ and $Q^2$ values of all models were over 0.85 (Fig. 1).

S-plot was used to elucidate compounds that were important for the classification based on covariance and loading profiling correlation of each metabolite. To limit the number of selected metabolites, absolute values of *p* and *p(corr)* were set to be greater than 0.2 and 0.5 (25). Sorbitol had higher *p* and *p(corr)* values then other compounds (Table S4). We then calculated maximum and minimum peak height and standard deviation of the selected compounds (Table S5). Peak height of each selected compound was significantly different between the groups, as evaluated by *t*-test. These observations indicated that S-plot-selected compounds are candidate bio-markers. To use these markers as thresholds in practical and industrial applications, however, the range of peak heights of the marker should not overlap with that in other groups. A box plot comparison of the range of peak heights of candidate markers among sample groups is shown in Fig. 2. Citric acid, gluconic acid, glucose, malic acid, proline and sorbitol were illustrated because these were selected from S-plot of OPLS-DA.

Peak heights of these compounds were normalized to that of ribitol, an internal standard. The result showed that proline and sorbitol could be used as single marker to discriminate *A. sinensis* from other species. This result was consistent with S-plot results, where sorbitol showed the closest $|p(corr)|$ to 1 of all compounds (Table S4). We then attempted to identify single-compound bio-markers for discrimination of *A. acutiloba* and *A. acutiloba* var. *sugiyamae*. However, the ranges of glucose_1 (Fig. 3A) or 4-aminobutyric acid (Fig. 3B) overlapped, and these two species could not be discriminated using a single-marker compound. To overcome this problem, we attempted to use two-compound bio-markers. Based on the S-plot, 4-aminoabutyric acid and glucose_1 appeared on the left and right side of the graph, respectively. Therefore, these two compounds contributed in discrimination of *A. acutiloba* and *A. acutiloba* var. *sugiyamae*. Overlap of box plots that are based on the range of peak height ratio between *A. acutiloba* and *A. acutiloba* var. *sugiyamae* could be resolved (Fig. 3). Hence, a two-compound bio-marker was used to successfully discriminate *A. acutiloba* from *A. acutiloba* var. *sugiyamae*.

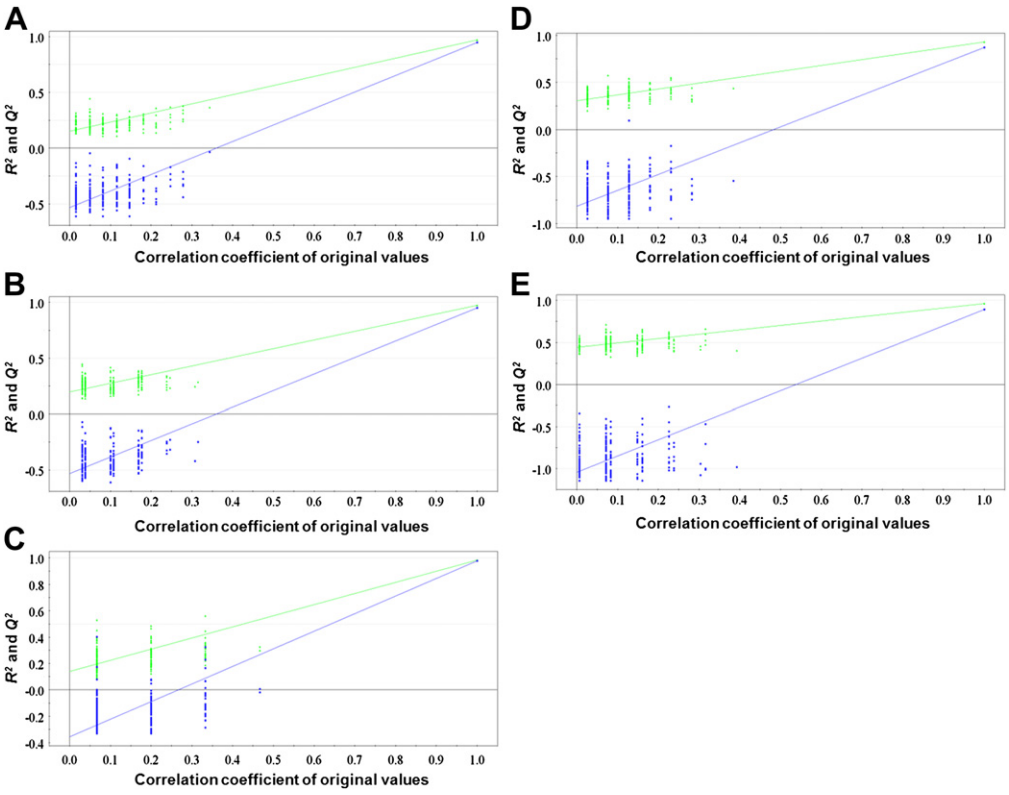


FIG. 1. Results from the permutation tests, which were carried out with 200 random permutations in PLS-DA models from GC–MS data. (A) 4 vs 3, 2, 1 model, (B) 4 vs 2, 1 model, (C) 4 vs 3, (D) 3 vs 2, 1 and (E) 2 vs 1.

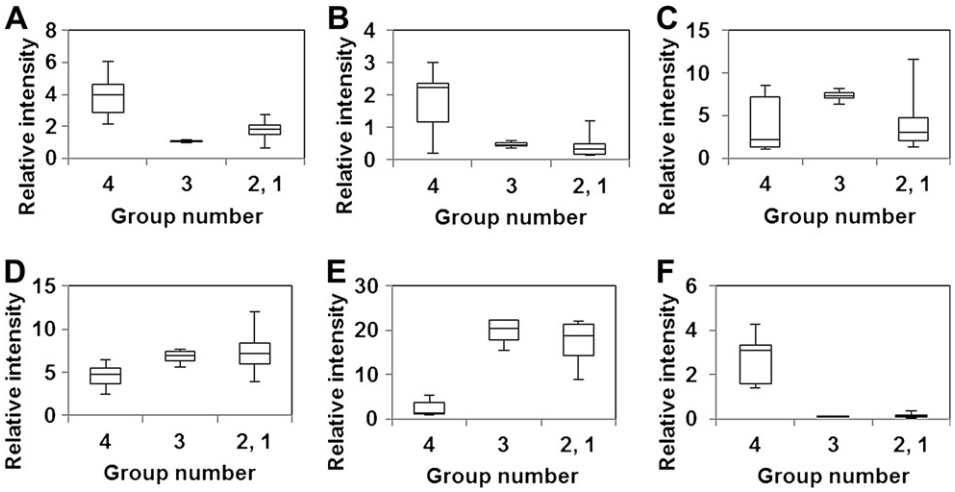


FIG. 2. Box plots of compounds selected by S-plot of OPLS-DA using relative GC–MS intensities to internal standard, ribitol: (A) citric acid, (B) gluconic acid, (C) glucose_1, (D) malic acid, (E) proline and (F) sorbitol.

Additionally, we investigated markers for discrimination of *A. acutiloba* samples based on agricultural origin using the ratio of peak heights selected by the S-plot (Table S5). Although there are several compounds that showed significant differences between the samples as evaluated by *t*-test, we could not achieve a separation without overlapping between samples in the box plot using two-compound bio-markers. Nevertheless, differences of markers that we identified using GC–MS profiles from *A. Radix* samples which contained proline/citric acid were significant by *t*-test based on species and cultivation area. In order to discriminate Angelica Radix samples based on cultivation area, the ratio between proline and citric acid ($p = -0.184$, $p(corr) = -0.535$) which was only peak $|p| > 0.15$ and $|p(corr)| = 0.5$ (citric acid was not used in the manuscript) were used to overcome an overlap between samples cultivated in China and Japan. Although the thresholds ($|p| > 0.2$ and $|p(corr)| = 0.5$) were appropriate to select markers in this method, unfortunately, an overlap was not improved as could be seen in Fig. S3. Discrimination of Angelica Radix based on cultivation area might be possible with the use of other analytical instrument such as inductively-coupled plasma mass spectrometry (ICP-MS). ICP-MS may be a suitable instrument to discriminate differences of cultivation area since it can analyze metal elements which directly have relationship with soil composition. However, if we take into consideration of many factors involved in quality control, the combination of GC–MS and GC–FID is a more appropriate technique because it can capture several factors such as species, cultivation area and processing methods and sensory quality. For the purpose of quality control in the market, the use of instruments that can assess multiple variables simultaneously would be desirable.

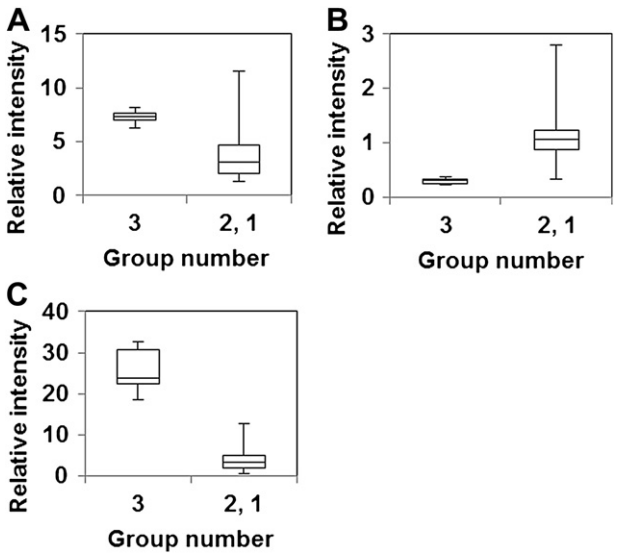


FIG. 3. Box plots of group 3 vs 1, 2 using relative GC–MS intensities: (A) glucose_1, (B) 4-aminobutyric acid, and (C) glucose_1/4-aminobutyric acid.

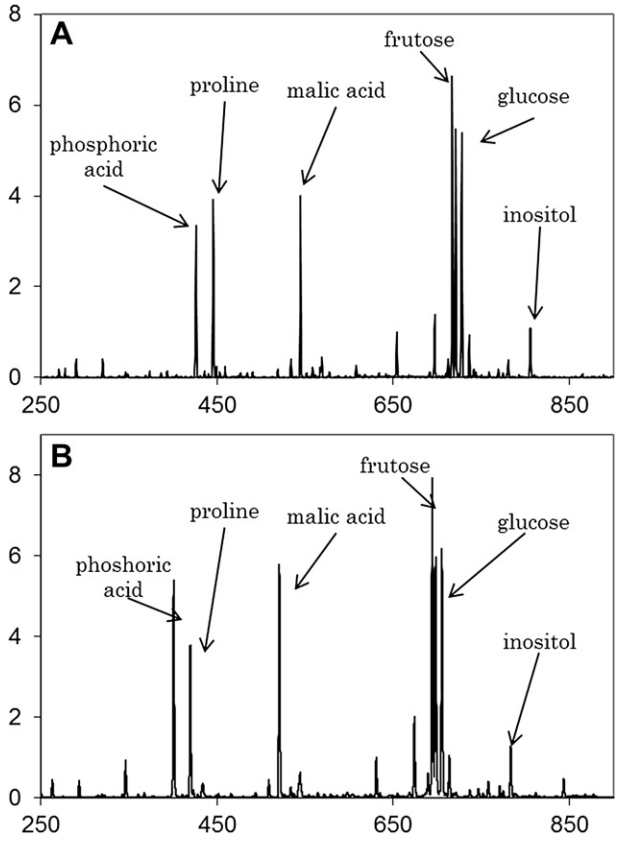


FIG. 4. Representative GC spectrum of *Angelicae Radix* (sample no. 1): (A) GC–MS spectra, (B) GC–FID spectra.

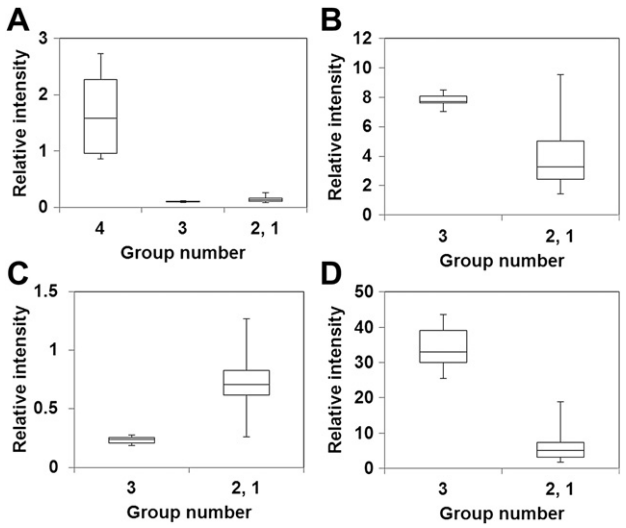


FIG. 5. Box plots of compounds selected by GC–MS data using relative GC–FID intensities to internal standard, ribitol: (A) sorbitol, (B) glucose_1, (C) 4-aminobutyric acid and (D) glucose_1/4-aminobutyric acid.

**Validation by GC–FID analysis** To make our results more applicable for use in industry, analysis by a simpler and more cost-effective instrument GC–FID was performed. We subjected the markers selected from GC–MS data – sorbitol, glucose and 4-aminobutyric acid to GC–FID and found that all markers could be used in GC–FID system. The chromatograms of GC–FID and GC–MS were shown in Fig. 4. All markers were checked by standard compounds. Retention times of sorbitol, glucose and 4-aminoglutamic acid were 705.0, 688.2 and 528.0 s, respectively. In Fig. 5, box plots of relative intensities obtained from GC–FID analyses were shown for validation. In Table S6, differences of all markers were significant by *t*-test. We proved that intensities of the same marker by GC–FID analysis were able to identify species and cultivation area of *A. Radix*.

In conclusion, we successfully discriminated species and cultivation area by GC–MS data and OPLS-DA. GC–MS intensities of markers selected by S-plot were statistically significant by *t*-test. Double-compound marker was efficient to increase relative intensity separation in box plot. In addition, the markers were applicable to a more conventional GC–FID analysis. The method shown in this study will be particularly useful for selection of markers to identify quality factors in commercial evaluation of not only other herbal medicines but also food products.

Supplementary data to this article can be found online at doi:10.1016/j.jbiosc.2012.03.022.

## ACKNOWLEDGMENTS

We would like to thank T. Mae of Maechu Co. Ltd. (Nara, Japan) and S. Sakamoto of Sakamotoshige shop for the sample information.

## References

1. **Fukuda, K., Murata, K., Matsuda, H., and Tani, T.:** History and the present of cultivation and production of Angelica root, Yamato-Toki, in Japan, Yakushigaku Zasshi, **44**, 10–17 (2009).
2. **Katoh, A. and Ninomiya, Y.:** Relationship between content of pharmacological components and grade of Japanese Angelica radixes, J. Ethnopharmacol., **130**, 35–42 (2010).
3. **Zhao, K. J., Dong, T. T., Tu, P. F., Song, Z. H., Lo, C. K., and Tsim, K. W.:** Molecular genetic and chemical assessment of radix Angelica (Danggui) in China, J. Agric. Food Chem., **51**, 2576–2583 (2003).
4. **Watanabe, A., Araki, S., Kobari, S., Sudo, H., Tchida, T., Uno, T., Kosaka, N., Shimomura, K., Yamazaki, M., and Saito, K.:** In vitro propagation, restriction fragment length polymorphism, and random amplified polymorphic DNA analyses of Angelica plants, Plant Cell Rep., **18**, 187–192 (1998).
5. **Murata, J., Oki, H., Kakutani, K., and Hashimoto, T.:** Identification of DNA polymorphisms in *Angelica acutiloba*, Plant Biotechnol., **25**, 157–163 (2008).
6. **Heubl, G.:** New aspects of DNA-based authentication of Chinese medicinal plants by molecular biological techniques, Planta Med., **76**, 1963–1974 (2010).
7. **Yip, P. Y., Chau, C. H., Mak, C. Y., and Kwan, H. S.:** DNA methods for identification of Chinese medicinal materials, Chin. Med., **2**, 9 (2007).
8. **Jones, C. J., Edwards, K. J., Castaglione, S., Winfield, M. O., Sala, F., Van de wiel, C., Bredemeijer, G., Vosman, B., Matthes, M., Daly, A., and other 12 authors:** Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories, Mol. Breed., **3**, 381–390 (1997).
9. **Piao, X. L., Park, J. H., Cui, J., Kim, D. H., and Yoo, H. H.:** Development of gas chromatographic/mass spectrometry-pattern recognition method for the quality control of Korean Angelica, J. Pharm. Biomed. Anal., **44**, 1163–1167 (2007).
10. **Yi, T., Leung, K. S., Lu, G. H., and Zhang, H.:** Comparative analysis of *Ligusticum chuanxiong* and related umbelliferous medicinal plants by high performance liquid chromatography–electrospray ionization mass spectrometry, Planta Med., **73**, 392–398 (2007).
11. **Cho, S. K., Abd El-Aty, A. M., Choi, J. H., Kim, M. R., and Shim, J. H.:** Optimized conditions for the extraction of secondary volatile metabolites in Angelica roots by accelerated solvent extraction, J. Pharm. Biomed. Anal., **44**, 1154–1158 (2007).
12. **Anetai, M., Aoyagi, T., Hayashi, H., and Furuki, M.:** Preparation and chemical evaluation of Angelicae Radix produced in Hokkaido. Part III. Time course study of dilute ethanol-soluble extract and sucrose contents, Rep. Hokkaido Inst. Public Health, **50**, 6–10 (2000) (in Japanese).
13. **Tarachiwin, L., Katoh, A., Ute, K., and Fukusaki, E.:** Quality evaluation of *Angelica acutiloba* Kitagawa roots by $^{1}$H NMR-based metabolic fingerprinting, J. Pharm. Biomed. Anal., **48**, 42–48 (2008).
14. **Tianniam, S., Tarachiwin, L., Bamba, T., Kobayashi, A., and Fukusaki, E.:** Metabolic profiling of *Angelica acutiloba* roots utilizing gas chromatography-time-of-flight-mass spectrometry for quality assessment based on cultivation area and cultivar via multivariate pattern recognition, J. Biosci. Bioeng., **105**, 655–659 (2008).
15. **Tianniam, S., Bamba, T., and Fukusaki, E.:** Pyrolysis GC-MS-based metabolite fingerprinting for quality evaluation of commercial *Angelica acutiloba* roots, J. Biosci. Bioeng. **109**, 89–93 (2010).
16. **Tianniam, S., Bamba, T., and Fukusaki, E.:** Non-targeted metabolite fingerprinting of oriental folk medicine *Angelica acutiloba* roots by ultra performance liquid chromatography time-of-flight mass spectrometry, J. Sep. Sci., **32**, 2233–2244 (2009).
17. **Kim, E. J., Kwon, J., Park, S. H., Park, C., Seo, Y. B., Shin, H. K., Kim, H. K., Lee, K. S., Choi, S. Y., Ryu, do, H., and Hwang, G. C.:** Metabolite profiling of *Angelica gigas* from different geographical origins using $^{1}$H NMR and UPLC-MS analyses, J. Agric. Food Chem., **59**, 8806–8815 (2011).
18. **Jumtee, K., Bamba, T., and Fukusaki, E.:** Fast GC–FID based metabolic fingerprinting of Japanese green tea leaf for its quality ranking prediction, J. Sep. Sci., **32**, 2296–2304 (2009).
19. **Lin, H., Zhuang, Q., and Fu, H.:** Habitat niche-fitness and Radix yield prediction models for *Angelica sinensis* cultivated in the Alpine area of the southeastern region of Gansu province, China, Plant Prod. Sci., **11**, 42–58 (2008).
20. **Tugawa, H., Tsujimoto, Y., Arita, M., Bamba, T., and Fukusaki, E.:** GC/MS based metabolomics: development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA), BMC Bioinformatics, **12**, 131 (2011).
21. **Verbruggen, N. and Hermans, C.:** Proline accumulation in plants: a review, Amino Acids, **35**, 753–759 (2008).
22. **Wiklund, S., Johansson, E., Sjostrom, L., Mellerowicz, E. J., Edlund, U., Shockcor, J. P., Gottfies, J., Moritz, T., and Trygg, J.:** Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models, Anal. Chem., **80**, 115–122 (2008).
23. **Wold, S., Sjostrom, M., and Eriksson, L.:** PLS-regression: a basic tool of chemometrics, Chemom. Intell. Lab. Syst., **58**, 109–130 (2001).
24. **Eriksson, L., Johansson, E., Kettaneh-Wold, N., and Wold, S.:** Muti- and megavariate data analysis principles and applications, p. 107. Umetrics AB, Umea, Sweden (2001).
25. **Llorach, R., Garrido, I., Urpi-Sarda, M., Tulipani, S., and Andres-Lacueva, C.:** Metabolomics study of human urinary metabolome modifications after intake of almond (*Prunus dulcis* (Mill.) D. A. Webb) skin polyphenols research articles, J. Proteome Res., **9**, 5859–5867 (2010).